Book Reviews

Reinforcement Learning: An Introduction—R. S. Sutton, and A. G. Barto (Cambridge, MA: MIT Press, 1998, 322 pp., hardcover, ISBN 0-262-19398-1) *Reviewed by O. Barana*

Overview of the volume

In the research area of artificial intelligence (AI) a branch that is becoming more and more important is reinforcement learning (RL). RL can be defined as learning how to map situations into actions interacting with the environment, so as to maximize a reward.

Written by two pioneers in this field, this book aims at supplying the basic RL ideas and algorithms. Even if the main point of view is the AI and engineering perspective, the Sutton–Barto book was designed to be accessible to readers of different disciplines. It turns out that the level of mathematical knowledge required to understand the material is not too deep and requires familiarity only with elementary notions of probability.

The book has been divided into three parts. The Problem (three chapters) is the introductory part devoted to the problem description. Elementary Solution Methods (three chapters) describes the most important elementary solution methods in authors' opinion: dynamic programming (DP), simple Monte Carlo (MC) methods, and temporal-difference (TD) learning. A Unified View (five chapters) concerns a generalization of the previous methods, gives a unified view of RL, and provides some examples of real RL applications.

Each chapter has many examples and exercises. Sections and exercises marked with a star (*) can be skipped during a first reading. At the end of every chapter there is a very interesting section dedicated to bibliographical and historical remarks.

Chapter 1: Introduction

The first chapter, an introduction to RL, emphasizes its main characteristics: interaction with environment, goal-directed learning, rewards, trade-off between exploration and exploitation. The four subelements into which a RL system can be subdivided are well explained: a policy, a reward function, a value function and, optionally, an environment model.

Chapter 2: Evaluative Feedback

One of the main characteristics of RL is that it uses training information that evaluates actions. This is well exemplified by the n-armed bandit problem described in this chapter, which in addition introduces the basic learning methods that will be used in the rest of the book. Concepts like action estimation, update rule, and greedy action selection become familiar.

Chapter 3: The Reinforcement Learning Problem

This is certainly the most important chapter of the first part, since it describes the RL problem and gives its mathematical formalization. For simplicity, the authors expressly consider discrete time steps and do not extend the treatment to the continuous time case. The Markov decision processes (MDPs), i.e., the RL tasks satisfying the Markov property, are defined: the finite MDPs (with finite sets of states and actions) represent the core of the processes on which RL theory is based. At the end of the chapter, the descriptions of the important state-value and action-value functions are given and the issue of the optimal policy is introduced.

The reviewer is with the Commissariat à l'énergie atomique, Département de recherches sur la fusion contrôlée, Centre de Cadarache, 13108 St. Paul lez Durance Cedex, France (e-mail: barana@drfc.cad.cea.fr).

Digital Object Identifier 10.1109/TNN.2004.842673

Chapter 4: Dynamic Programming

Describing the DP, Chapter 4 begins the part devoted to the three fundamental elementary solution methods of the RL problem. Under the name of DP a series of algorithms used to compute an optimal policy, given the perfect knowledge of the environment model, can be found. The authors start dealing with the so-called iterative policy evaluation (prediction phase) which can lead to the determination of the optimal policy—the continuous process of interaction between policy evaluation and policy improvements, in order to determine the optimal policy, is called generalized policy iteration (GPI) and is a constant throughout the book. Considerations on DP advantages, disadvantages, and efficiency close this chapter.

Chapter 5: Monte Carlo Methods

MC methods present two fundamental differences with respect to DP: a) they are based on experience (online or simulated) and therefore do not require knowledge of the environment's model; b) the computed estimates do not depend on other estimates. MC methods are needed to solve RL problems based on averaging sample returns, and the updates are not incremented step-by-step, but episod-per-episod. In this chapter, the concept of GPI is well analyzed and emphasis is given to the need of an adequate exploration, which can be equally achieved with both on-policy and off-policy methods.

Chapter 6: Temporal-Difference Learning

In Chapter 6, the authors discuss the temporal-difference (TD) learning. This method has the characteristics of both DP and MC learning. Like DP, TD learning updates its estimates partly relying on other estimates (this is called bootstrapping). Like MC, TD learning is based on experience, without knowledge of the environment's model. This chapter compares mainly TD with MC learning. GPI is treated as well: in this phase the Sarsa and Q-learning methods acquire a particular relevance. Finally, it is interesting to note that the methods explained in this chapter are the most exploited nowadays to solve RL problems.

Chapter 7: Eligibility Traces

Chapter 7 is the first of the third part, devoted to a unified view of the three basic methods used to solve RL problems presented in the book. In particular, it is explained how to link MC and TD methods. This is achieved with the use of the eligibility traces (ET), considered by the authors as one of the basic mechanism of RL. ET methods can be used in GPI, are valid for non-Markov tasks and are also preferable in online applications, since they learn in a faster way than one-step methods like the simple TD.

Chapter 8: Generalization and Function Approximation

Chapter 8 is very interesting because it combines the RL world with that of supervised learning, disclosing several fascinating horizons for RL applications. By means of the interaction with supervised approximation methods, the authors explain how generalizing properties of RL increase enormously. Particularly, the gradient-descent linear method is described in details (even though it is said that the nonlinear method of backpropagation neural networks is good as well). After having treated prediction and GPI like in the previous chapters, the final considerations are made between function approximation, bootstrapping and distinction between on-policy and off-policy.

Chapter 9: Planning and Learning

This chapter provides a unified view of RL methods requiring a model (called planning methods) and of those that do not require a model (called learning methods). Other reflections on different RL

0018-9448/$20.00 © 2005 IEEE

methods derive from considerations over different type of backups (the updates of the states). The chapter is closed by a "trespass" in the AI field: implications in exploiting heuristic search with planning methods are analyzed.

Chapter 10: Dimensions of Reinforcement Learning

In chapter 10, the authors present the final considerations on RL, highlighting how, despite the versatility and power of this "philosophy," a lot remains to be studied in the RL research area (for example the extension to non-MDPs). They show how all three RL methods share three elements (estimate of a value function, backup operations, processes of GPI), and give a description of the space of RL methods.

Chapter 11: Case Studies

This last chapter deals with interesting applications of RL techniques to the real world, illustrating how many of them have a potential economic importance.

Summary

Written to be a reference book to be used in university courses, the Sutton–Barto book is nevertheless something more, since in my opinion it is not only an introduction to the RL issue (as the title says), but also supplies the basis for development of RL applications in several fields of the real world. Easy to understand but sufficiently rigorous at the same time, this book is the right candidate to be one of the milestones in the RL research and in AI in general. It will be surely appreciated also by those who have a keen interest in supervised learning, who will see in the content of this book a way to extend and improve their applications.

Evolving Connectionist Systems: Methods and Applications in Bioinformatics, Brain Study and Intelligent Machines—N. Kasabov (New York: Springer-Verlag, 2003, ISBN 1852334002) *Reviewed by G. S. Ng*

The main purpose of the book is to present a generic computational model and techniques that can be used for the development of evolving modeling systems. The book consists of two parts. The first part presents generic methods and techniques for evolving connectionist systems. The second part presents applications.

On the whole, it is an excellent book and will be of use to a large number of researchers who are interested in evolving systems. Of particular interest to the readers are those chapters in Part II where related applications are presented. These applications cover most of the applications in artificial intelligence such as bioinformatics, speech recognition, and image classification.

The presentation order, notation, and glossary are clear and well organized for quick reference. This is precisely what we seek in a book, what we would want on a shelf close at hand. The huge number of references makes it difficult for the reader to identify the essential ones.

The book covers, to some extent, genetic and neuronal levels of brain modeling. Currently, this is a hot and challenging topic among researchers. Another excellent topic is, in Chapter 7, devoted to the multiagent implementation framework for building evolving connectionist machines that integrate several evolving connectionist systems together to solve a given task. The modules, as well as their internal structures, evolve in time.

Readers of the book may be slightly disappointed by the coverage of applications chapters. After reading this part I have the impression that various applications of evolving connectionist system exist, but how exactly they are used is not so obvious.

The book layout is clear and uniform throughout. However, the notation reveals some nonuniformities. For example, superscript t is used to represent t-th iteration (or tth presentation of data pair) in one place, but later in the text tth presentation is denoted as subscript.

The readers of this book will find the clear and comprehensive topics and list of references of the greatest use. Overall, this book is recommended for readers who would like to enter the subject of evolving systems.

Data Mining. Multimedia, Soft Computing, and Bioinformatics—S. Mitra and T. Acharya (New York: Wiley, 2003, pp. 401 + xiv, ISBN 0-471-46054-0) *Reviewed by W. Pedrycz*

There is no doubt that with the growing role of the Internet, omnipresent visibility of inherently heterogeneous information typical for multimedia, vast mountains of data generated in a twinkle of an eye, an efficient, user-centric and successful search for the meaningful relationships in data becomes of a paramount importance and genuine necessity.

The book by Mitra and Acharya lies in the center of the recent developments in data mining. As the title itself stipulates, the authors have positioned data mining in the framework of multimedia, soft computing, and bioinformatics. The title is highly promising—so is the book. The authors, who are authorities in the area, show us how to navigate across the oceans of data, avoid the reefs of memorization and curses of dimensionality. Multimedia information is overwhelming today. It is right to stress that soft computing (or computational intelligence as these two names are often used interchangeably) plays a pivotal role as a conceptual and algorithmic platform of casting data processing in a user-centric environment guided by a transparent logic processing of information granules.

As in case of any book, it is instructive to review its content. This may serve two purposes. First, we learn about the breadth of coverage of the area. Second, it tells us about the linkages between the dominant topics of the book and helps us learn about the architecture of the domain. Chapter 1 is a concise introduction to data mining being viewed here not only as a rapidly growing area of intensive research endeavors but also a territory of numerous case studies and real-world applications. The notions like information retrieval, text and image mining, set up the stage for the contemporary agenda of data mining. Chapter 2 forms a comprehensive suite of generic prerequisites of soft computing. It starts with a descriptive definition of soft computing by emphasizing how different contributing technologies (fuzzy sets, neural networks, rough sets, and evolutionary optimization) build a highly collaborative design environment of intelligent systems and intelligent data analysis. Chapter 3 is devoted to a central topic of multimedia that is data compression. Compression permeates various facets image processing (as a matter of fact, it is common that images are highly redundant in the sense of information captured by individual pixels or regions). The chapter reports also on the classic methods of text compression such as Lempel–Ziv (usually coming under familiar names of LZ77, LZ78, LZW, and alike). The material of string matching (Chapter 4) is a classic prerequisite to any sound research in bioinformatics with strings and their comparison being a bread-and-butter of the research going there. Classification models in data mining are covered in Chapter 5.

The reviewer is with the School of Computer Engineering, Nanyang Technological University, Singapore 639798.

Digital Object Identifier 10.1109/TNN.2004.842676

The reviewer is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada (e-mail: pedrycz@ee.ualberta.ca).

Digital Object Identifier 10.1109/TNN.2004.842675

0018-9448/$20.00 © 2005 IEEE